Supporting Information

Sharma et al. 10.1073/pnas.1409543112

SI Text

Exchange of Modeled Water Molecules with the Crystallographic Waters. The crystallographic water molecules present in the hydrophilic region above the propionates diffuse into the active-site region, and engage in water wire formation (Fig. S1). Out of four modeled water molecules (see *Materials and Methods* in the main text) in the nonpolar cavity (wat#1–4, Fig. S1), one water molecule (wat#1) escaped the cavity region as early as around 10 ns. Out of the three remaining water molecules which participated in the water wire formation (Fig. 2/PUMP), wat#3 escaped the cavity at around 55 ns. The escape route of water molecules in these simulations is found to be similar to what has been observed previously, i.e., close to the Arg438–Dprp (heme a_3)–Trp126 region (1, 2).

Except for the crystallographic water molecule, which diffused into the cavity and participated in water wire formation for around 10 ns (wat#5, Fig. S1), other crystallographic water molecules (wat#6 and wat#7) diffused into the cavity only at a later stage (30–40 ns), and participated in water wire formation. The entry path is found to be similar to the exit path described previously, i.e., the Arg438–Dprp (heme a_3)–Trp126 region (1, 2). The water exchange takes place on a time scale of tens of nanoseconds, similar to exit rates of water molecules proposed previously (2).

Grand Canonical Monte Carlo Simulations for Water Occupancy. The grand canonical Monte Carlo (MC) method has been successfully used in the past to identify potential locations of water molecules in the interior of proteins (3). We applied a similar methodology using CHARMM (4) to ascertain the level of hydration in the nonpolar cavity of C*c*O. We used two different values for the excess chemical potential of water ($\mu_{ex} = -5.8$ and -6.8 kcal/mol). We observed that no water molecules are predicted in the cavity for up to 5 million MC steps in the case when Glu242 points down, as in the crystal structure. However, when simulations were performed with Glu242 in an up conformation, which is the relevant case for the current work, 3 and 4 water molecules were predicted at $\mu_{ex} = -6.8$ and -5.8 kcal/mol, respectively. This is in agreement with previous estimates (5), and supports the modeling of four water molecules in the current work.

Free-Energy Calculations. We performed umbrella sampling simulations in the $\mathbf{P_M}$' state to estimate the free-energy difference between the CHEM and PUMP configurations, and energetics associated with the formation of the CHEM configuration. We introduced two collective variables to describe the collective motion of the water molecules during the CHEM ↔ PUMP transition (Fig. 2). The first collective variable (x) defines the average of three O–O distances between the neighboring water molecules forming a four-membered continuous water chain between Dprp of heme a_3 and the BNC (CHEM configuration, Fig. 2). The second collective variable (y) defines the distance between the O atom of the water molecule closest to the Dprp and the protonated O atom of Glu242 (CHEM configuration, Fig. 2). To generate initial conformations for the umbrella sampling windows and to test our collective variables in realizing the CHEM ↔ PUMP transition, we first performed multiple short (1–5-ns) steered molecular dynamics (SMD) simulations with different spring constants and speeds, in which the value of the first collective variable is restrained within a hydrogen-bonding range ($x \le x_0 = 3$ Å) with a half-harmonic potential, $U_1 = \begin{cases} 0, & x \le x_0 \\ \frac{1}{2}k_1(x-x_0)^2, & x > x_0 \end{cases}$, whereas the value of the second collective variable is increased linearly in time in terms of $U_2(t) = 1/2k_2(y-(y_0+vt))^2$. SMD simulations resulted in a smooth and continuous CHEM ↔ PUMP transition and allowed us to capture intermediate states. We obtained initial conformations from one of the SMD simulations for a total of 14 umbrella sampling windows, which are selected and simulated progressively to improve sampling. The same two collective variables (x and y) used in SMD simulations were biased in umbrella windows, first with the half-harmonic potential, U_1, with $k_1 = 2.5$ kcal·mol^{-1}·Å^{-2} and $x_0 = 3$ Å, and second with a harmonic potential, U_2, with $k_2 = 3$ kcal·mol^{-1}·Å^{-2} for 11 windows and $k_2 = 10$ kcal·mol^{-1}·Å^{-2} for 3 windows (to improve sampling of the CHEM configuration) with harmonic potential centers in [5.4 Å, 11.4 Å]. Each window was simulated for 2–7 ns, with longer simulations performed for windows with higher autocorrelation times to obtain better sampling. The total simulation time is 50.1 ns. Samples within five autocorrelation times of each simulation were discarded from the beginning before any analysis. Sample reweighting using a nonparametric variation of weighted histogram analysis method (WHAM) (6), analogous to the multistate Bennet acceptance ratio method (7), and error analysis using Bayesian block bootstrapping (8) were performed as implemented in ref. 9. For Bayesian block bootstrapping, each window was divided into three blocks, each with a width of at least five autocorrelation times of the window it belongs to, and 100 bootstrap samples were generated. The weights obtained were used to estimate all of the free-energy differences and to reconstruct the potential of mean force (PMF) profiles.

The free energy for formation of the CHEM configuration in the $\mathbf{P_M}$' state is estimated to be 3.14 ± 0.33 kcal/mol with respect to no CHEM configuration. This result, in agreement with the unbiased simulations (Table 2), suggests that the CHEM configuration is not favorable in the $\mathbf{P_M}$' state. Here, we also present a mechanistic picture of the formation of the CHEM configuration. Fig. S2 shows PMF as a function of distance between the O atom of the –OH ligand of Cu_B and the O atom of the closest water molecule. A distance of ∼6 Å corresponds to the location of E242. The red curve is the PMF obtained using whole data, whereas the green and the blue curves are based on the portion of the data in which the PUMP and the CHEM configurations form, respectively. Here, we divide the formation of the CHEM configuration (in the $\mathbf{P_M}$' state) into two steps: (*i*) entry of at least one water molecule into the region between E242 and Cu_B, and (*ii*) reorganization of the water molecules into a continuous hydrogen-bonded chain. Fig. S2 shows that the formation of the CHEM configuration is associated with step (*i*), because the CHEM configuration (blue curve) is restricted to (2, 4) of the x axis. The main free-energy barrier associated with step (*i*) is 2.63 ± 0.15 kcal/mol at a distance of 4 Å away from the Cu_B O atom. On the other hand, the shift between the blue and the red curves (∼1.4 kcal/mol) corresponds to step (*ii*), that is, the reorganization of the water molecules into a continuous hydrogen bonded chain when there is at least one water molecule hydrogen bonded to Cu_B.

Effect of Charge Variation on the Water-Gate Mechanism. To assess how robust or sensitive the main conclusions of our work are on the charge parameterization (force field) used in the simulations, we also performed *ca*. 15-ns-long MD simulations in the $\mathbf{P_M}$' and $\mathbf{P_R}$ states using the assisted model building with energy

refinement (AMBER) force field (10) charge set, as described in ref. 11. A rapid escape of water molecules from the nonpolar cavity was observed, which resulted in reduced populations of both PUMP and CHEM configurations compared with Table 2. Nevertheless, the results show that there is a 10-fold preference for the PUMP configuration (relative to CHEM) in the $\mathbf{P_M}$' state, thereby supporting the water-gate mechanism.

Resampling of the Energetics Data. To test the robustness of the energetics data presented in Table 2, we performed resampling of the CHEM configuration data obtained from $\mathbf{P_M}$', $\mathbf{F_H}$', and $\mathbf{F'_C}$ state simulations. We generated 1,000 replicates of the data (with or without replacement) and evaluated the SDs, which are as follows: $\mathbf{P_M}$' (3.8 ± 0.03 kcal/mol), $\mathbf{F_H}$' (3.4 ± 0.01 kcal/mol), and $\mathbf{F'_C}$ (1.0 ± 0.001 kcal/mol).

1. Wikström M, et al. (2005) Gating of proton and water transfer in the respiratory enzyme cytochrome *c* oxidase. *Proc Natl Acad Sci USA* 102(30):10478–10481.
2. Sugitani R, Stuchebrukhov AA (2009) Molecular dynamics simulation of water in cytochrome *c* oxidase reveals two water exit pathways and the mechanism of transport. *Biochim Biophys Acta* 1787(9):1140–1150.
3. Woo H-J, Dinner AR, Roux B (2004) Grand canonical Monte Carlo simulations of water in protein environments. *J Chem Phys* 121(13):6392–6400.
4. Brooks BR, et al. (2009) CHARMM: The biomolecular simulation program. *J Comput Chem* 30(10):1545–1614.
5. Riistama S, et al. (1997) Bound water in the proton translocation mechanism of the haem-copper oxidases. *FEBS Lett* 414(2):275–280.
6. Bartels C (2000) Analyzing biased Monte Carlo and molecular dynamics simulations. *Chem Phys Lett* 331(5-6):446–454.
7. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129(12):124105.
8. Hub JS, de Groot BL, van der Spoel D (2010) g_wham—A free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J Chem Theory Comput* 6(12):3713–3720.
9. Moradi M, Tajkhorshid E (2014) Computational recipe for efficient description of large-scale conformational changes in biomolecular systems. *J Chem Theory Comput* 10(7):2866–2880.
10. Duan Y, et al. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* 24(16):1999–2012.
11. Johansson MP, Kaila VRI, Laakkonen L (2008) Charge parameterization of the metal centers in cytochrome *c* oxidase. *J Comput Chem* 29(5):753–767.

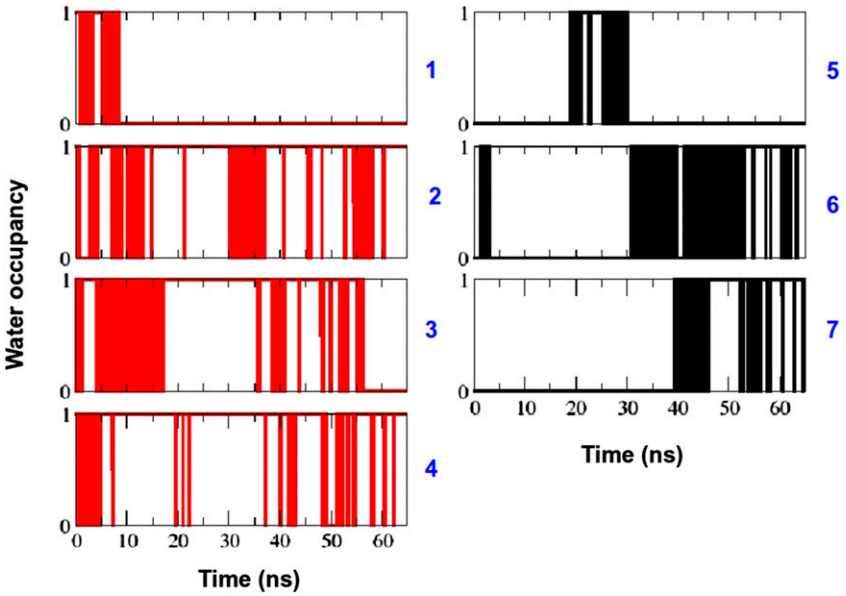

Fig. S1. Occupancy of water molecules in the active-site region. The occupancy of water molecules is calculated from the $\mathbf{P_M}$' state simulation. Water molecules are considered to be present in the region (*y* axis = 1) if the distance (*dist*) between the oxygen atom of the water molecule is <10 Å from protonated Glu242, from the Dprp of the high-spin heme and from the OH ligand of Cu_B; otherwise they are considered to be absent (*y* axis = 0). Red traces correspond to four water molecules that were modeled in the beginning of the simulation (wat#1–4; see *Materials and Methods* in the main text), whereas the black traces correspond to crystallographic water molecules (wat#5–7) initially present in the hydrophilic domain above the propionates.

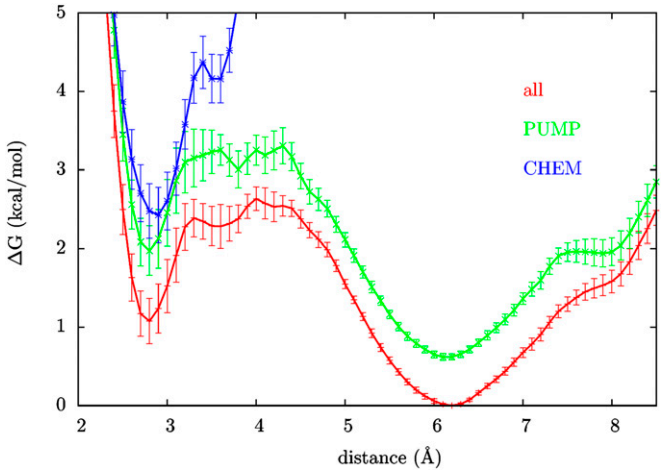

Fig. S2. PMF profiles as a function of distance between the Cu_B O atom and the O atom of the closest water molecule. Red, green, and blue lines correspond to PMF calculated using the whole data, only the PUMP state configurations, and only the CHEM state configurations, respectively.

Table S1. Orientation of water chains in the nonpolar cavity above Glu242 with different numbers of water molecules above and below Glu242

Redox state	PUMP*, %	CHEM*, %
$P_M'^{\dagger}$	23	0.01
$P_R^{\dagger}$	25	20
$P_M'^{\ddagger}$	36	3
$P_R^{\ddagger}$	6	32

*PUMP denotes a complete hydrogen-bonded water wire from Glu242 to the Dprp of the high-spin heme, CHEM denotes a corresponding water wire from Glu242 to the binuclear center (see hydrogen bonding criteria in *Materials and Methods*). The values are the percentage of frames with PUMP or CHEM configuration out of all frames of the simulation.

†Three additional water molecules modeled below Glu242 and five in the cavity above Glu242.

‡Three additional water molecules modeled below Glu242 (four in the cavity above Glu242, as in Table 2).

Table S2. Orientation of water chains in the nonpolar cavity in different redox states determined from in vacuo simulations

Redox state	PUMP*	CHEM*
F_H'	16	0.001/7.0†
$F_{H,R}$	5	35
F'	19	0/NA†
F_R	20	28
F'_C	37	32/0.7†

*Percentage of frames in which a complete water wire from Glu242 to the Dprp of high-spin heme (PUMP) or to the BNC (CHEM) is observed (hydrogen-bonding criteria as discussed in *Materials and Methods*). A total of 100,000 frames were analyzed. NA stands for not available.

†Energy cost [being described as $-k_BT$ ln P(CHEM) in units of kcal/mol] associated with the formation of the CHEM configuration in the states when heme *a* is reduced.